8pp

OPPT-2002-0029-0013

3. ICCVAM-NICEATM - In Vitro ER/AR Assays

→ 3.1 Estrogen Receptor Binding Assays – Executive Summary

3.2 Androgen Receptor Binding Assays - Executive Summary

3.3 Estrogen Receptor Transcriptional Activation Assays - Executive Summary

3.4 Androgen Receptor Transcriptional Activation Assays - Executive Summary

For more information and the complete reports on these assays, please go to the web page at
http://iccvam.niehs.nih.gov/methods/endodocs/ed_brd.htm

Current Status of Test Methods for Detecting Endocrine Disruptors: *In Vitro* Estrogen Receptor Binding Assays

Prepared by
The National Toxicology Program (NTP)
Interagency Center for the Evaluation of Alternative Toxicological Methods
(NICEATM)

National Institute of Environmental Health Sciences
(NIEHS)
P.O. Box 12233
Mail Drop: EC-17
Research Triangle Park, NC 27709

DRAFT

April 2002

EXECUTIVE SUMMARY

The objectives of this BRD are to: (1) provide comprehensive summaries of the published and publicly available unpublished data on the scientific basis and performance of *in vitro* assays used to test substances for their ability to bind to the estrogen receptor (ER); (2) assess the *in vitro* ER binding assays considered for their effectiveness in identifying endocrine-active substances; (3) identify and prioritize *in vitro* ER binding assays that might be considered for incorporation into future testing programs for validation; 4) develop minimum performance criteria by which to judge the effectiveness of proposed *in vitro* ER binding assays; and (5) generate a list of recommended substances to be used in validation efforts.

The data summarized in this BRD are based primarily on information obtained from the peer-reviewed scientific literature. An online literature search was conducted to retrieve records on publications reporting on the testing of substances for their endocrine disrupting effects *in vitro*. Of the 459 records obtained from the initial search, 260 of these citations contained information on ER binding. Data from 74 of these publications were included in this BRD. Some of the peer-reviewed publications that contained ER binding data were not abstracted for inclusion in this BRD because the studies lacked the appropriate details or contained data from unique procedures or substances that were not clearly identified.

Data were abstracted from 14 different ER-binding assays. These assays used ER derived from uterine cytosol from the mouse (MUC), rat (RUC), and rabbit (RBC); from MCF-7 cells and MCF-7 cytosol; from human cDNA clones of the two human ER isoforms, ERα and ERβ (hERα and hERβ). Fusion proteins in which glutathione (GST) was fused with the *def* domains of the human ERα (GST-hERα), and the ER from mice (GST-mER), chicken (GST-cER), anole (GST-aER), and rainbow trout (GST-rtER) was the basis for five assays. All the assays except one measured the competitive displacement of [^{3}H] 17β-estradiol from the ER. One assay, designated as hERα-FP, measured the displacement of a fluorescently-labeled estradiol by the test substances using fluorescent polarization (FP).

The majority of the 635 substances tested for *in vitro* ER binding could be classified into one of the following chemical classes: polychlorinated biphenyls, phenolic and non-phenolic steroids, triphenylethylenes, organochlorines, stilbene analogs, polycyclic aromatic hydrocarbons, phenols, indenes, bisphenols, and flavonoids. Only 48% of the substances could be assigned to a product class, the most common of which were pharmaceuticals, pesticides, dielectric fluids or their components, chemical intermediates, natural products (including several phytoestrogens), and plasticizers.

More than half the substances (377; 59%) were tested in the RUC assay, and 135 (21%) of the substances were tested in the next most frequently used assay, hERα. For five of the 14 assays (hERα-FP, RBC, rERβ, GST-mERαdef, GST-cERdef), published data were located on fewer than 50 substances per assay. Only 2.4% of the substances had been tested in 10 or more assays, and of these, only four (0.8%) had been tested in all 14 assays; in contrast, 397 (62.5%) of substances had been tested in one assay only.

The majority of the publications reported the data as IC_{50} values or relative binding affinities (RBA), that is, the ratio of the IC_{50} of the reference estrogen, 17β-estradiol, divided by the IC_{50} of the test substance x100.

Although a large number of substances have been tested in these *in vitro*-ER binding assays, relatively few have been tested more than once in the same assay or in multiple assays. Furthermore, because the primary focus of many of the studies reviewed in this BRD focused on understanding the mechanisms of ER binding and not at identifying substances with ER-binding activity, much of the published data are of limited value for the analysis of performance or reliability of these assays.

To assess comparative assay performance, a quantitative assessment was conducted using the available IC_{50} and RBA data after log normal transformation of the data to reduce possible skewness. In this analysis, only positive responses were considered (i.e., discordant positive and negative results for the same substance in the same assay were not taken into account). The quantitative assessment of the data showed that the effect of substances on the variation in RBA

and IC_{50} values was much greater than the effect of assay type, and that there were no significant differences in performance among the different *in vitro* ER binding assays. This quantitative assessment was limited by the lack of multiple test data within an assay for most of the substances, and by the lack of data across all assays for many substances.

A qualitative assessment of the IC_{50} data, which considered both negative and positive results, was performed also. This assessment considered whether RBA values (single or median) obtained for substances tested in each of 13 assays were within the same log range as the corresponding values obtained for the same substances when tested in the RUC assay. Based on this qualitative approach, the hERα, hERα-FP, hERβ, rERβ, GST-rtERdef, and MUC assays performed better than the RUC assay; the MCF-7 cytosol assay performed about as well as the RUC assay; and the remaining GST-ERdef assays, the MCF-7 cell assay, and the RBC assay did not perform as well as the RUC assay.

To assess assay reliability, a quantitative assessment was conducted using log normal transformed IC_{50} and RBA data. Again, only positive responses were considered. An analysis of the variances for the RBA values of 12 substances that had been tested in at least ten of the 14 *in vitro* ER binding assays suggested that there were no statistically significant differences in the reliability of the assays as performed by different laboratories. A comparison of the variability in RBA and IC_{50} values across assays, ignoring substance effects, suggested that the RUC and hERβ assays were the most consistent, and that the RBC assay was the least consistent among the 14 assays evaluated. An analysis of the variability in the IC_{50} for the reference control chemical, 17β-estradiol, indicated that the most consistent results were obtained with the hERα-FP assay, while the MUC, RUC, and hERα assays exhibited somewhat greater, but comparable, variances. The low variability associated with the hERα-FP assay, however, might be a reflection of the small number of laboratories that have reported IC_{50} values using this method.

Generally, the databases for all the *in vitro* ER binding assays considered in this BRD are too limited to draw any sound conclusions regarding their performance and reliability. However, based on general principles, recommendations were made in regard to the use of *in vitro* ER binding assays as a component of a Tier 1 endocrine disruptor screening battery

- Based on a consideration of such factors as relative performance, elimination of animal use, the use of the ER from the species of interest, and the use of alternatives to radioactive substances, the hERα, hERα-FP, and hERβ assays should have the highest priority for validation as screening assays for human health-related issues, while the GST-rtERdef assay might be preferred when screening for substances that pose an hazard to wildlife.
- In conducting future validation studies with these assays, the RUC assay should be used as the reference test method. The RUC assay is currently undergoing validation efforts sponsored by the U.S. EPA and the resulting performance and reliability information could be used to establish minimal performance standards for other assays.
- Formal validation studies should be conducted using appropriate substances covering the range of expected RBA values to adequately demonstrate the performance characteristics of the *in vitro* ER-binding assays recommended as possible screening assays.
- There is little information about the ER binding activity of metabolites of xenobiotics and it is not clear whether metabolic activation needs to be included in *in vitro* ER binding test methods used as screening assay. This issue should be considered prior to the implementation of future validation studies.

An important step towards acceptance of an *in vitro* ER binding assay into a regulatory screening program is production of high quality data. To achieve this goal, it is recommended that any future pre-validation and validation studies on *in vitro* ER binding assays be conducted with coded substances and in compliance with GLP guidelines. Ideally, if multiple laboratories are involved in the validation study, the substances should be obtained from a common source and distributed from a central location.

In conducting these validation studies, all of the original data and documentation supporting the validation of a test method must be carefully documented, and include detailed protocols under which the data were produced.

If an assay chosen for validation requires the use of animals, the studies should be conducted to minimize the number of animals used, and animal pain and distress. Adoption of one of the

assays using purified or semi-purified receptors, or glutathione fusion proteins would eliminate the use of animals for *in vitro* ER binding experiments.

Since there are no published guidelines for conducting *in vitro* ER binding studies, and no formal validation studies have been performed to assess the reliability or performance of ER binding assays, the U.S. EPA requested that minimum procedural standards based on a comparative evaluation of *in vitro* ER binding assays be provided. In addition it was requested that a recommended list of test substances be provided for use in validation studies,

The minimum procedural standards include methods for determining the K_d of the reference estrogen, methods for test substance preparation, the concentration range of the test substance (including the limit dose), the use of negative and positive controls, the number of replicates per test substance concentration, dose spacing, assay acceptance criteria, data analysis, evaluation and interpretation of results, minimal information to include in the test report, and the need for replicate studies.

Based on a RUC protocol provided by the U.S. EPA, a suggested general protocol for measuring ER binding using the RUC assay was developed as a potential resource for scientists interested in developing their own laboratory specific protocol. This general RUC protocol incorporated the recommended minimum procedural standards. Various aspects of the assay performance, including preparation of the ER, reagents and solutions, measurement of ER-binding, evaluation of the data, and test report guidance are described.

In the development of a list of reference substances for use in validation studies, consideration was given to the number of times the substance had been tested in the RUC assay, the median RBA value of the substance in the RUC assay, and the extent of concordance of the median RBA value in the RUC with values obtained for the same substance in other *in vitro* ER binding assays. The substances were then sorted according to their median RBA values, which ranged over seven orders of magnitude. Five substances were selected for each RBA category ($\geq$10, <10-1; <1-0.1; <0.01-0.1, <0.01-0.001; <0.001) and three for a negative category group. Weakly-binding substances (RBA values <0.001) were difficult to identify because they were

not consistently positive in tests within an assay or among different assays. Substances were classified as "negative" for ER binding based on the lack of a positive response in multiple assays when tested at concentrations of at least 1 mM. When possible, representatives of the most common classes of substances were included in each RBA category.

It is anticipated that this BRD and the guidance it provides will help to stimulate validation efforts for *in vitro* ER binding assays.